Association between long-term exposure to air pollution and specific causes of mortality in Scotland

Post-print of peer-reviewed article published by BMJ group.
Published article available at: http://dx.doi.org/10.1136/oemed-2011-100600

Cite as:
Yap, C., Beverland, I. J., Heal, M. R., Cohen, G. R., Robertson, C., Henderson, D. E. J., Ferguson, N. S., Hart, C. L., Morris, G., Agius, R. M., 2012. Association between long-term exposure to air pollution and specific causes of mortality in Scotland.
***Occupational & Environmental Medicine* 69, 916-924.**

Association between long-term exposure to air pollution and specific causes of mortality in Scotland

Christina Yap[1,2], Iain J Beverland[1*]; Mathew R Heal[3]; Geoffrey R Cohen[4]; Chris Robertson[5,6,7]; Deborah E J Henderson[1], Neil S Ferguson[1], Carole L Hart[8], George Morris[9] and Raymond M Agius[10]

[1]Department of Civil Engineering, University of Strathclyde, UK
[2]Current address: MRC Midland Hub for Trials Methodology Research, University of Birmingham, UK
[3]School of Chemistry, University of Edinburgh, Edinburgh, UK;
[4]Edinburgh, UK;
[5]Mathematics and Statistics, University of Strathclyde, Glasgow, UK;
[6]Health Protection Scotland, Glasgow, UK;
[7]International Prevention Research Institute, Lyon, France
[8]Institute of Health and Wellbeing, Public Health, University of Glasgow, Glasgow, UK;
[9]National Health Service in Scotland, Glasgow, UK
[10]Centre for Occupational and Environmental Health, The University of Manchester, Manchester, UK;

***Correspondence to:**
Dr Iain Beverland,
Senior Lecturer in Environmental Health
University of Strathclyde, Department of Civil Engineering,
John Anderson Building, 107 Rottenrow, Glasgow G4 0NG
Tel: 0141-548-3202
Fax: 0141-553-2066
Email: Iain.Beverland@strath.ac.uk

Abstract:

Objective: This study investigated the association between long-term exposure to black smoke (BS) air pollution and mortality in two related Scottish cohorts with 25 years of follow-up.

Methods: Risk factors for 15,311 participants in Renfrew/Paisley and 6680 participants in Collaborative cohorts were collected during 1970-1976. Exposure to BS during 1970-79 was estimated by inverse-distance weighted averages of observed concentrations at monitoring sites, and two spatial modelling approaches which included local air quality predictors (LAQP).

Results: Consistent BS-mortality associations (per 10 μg m^{-3} increment in 10-year average BS) were observed in the Renfrew/Paisley cohort using LAQP-based exposure models [all-cause mortality hazard ratio (HR) = 1.10 (95% confidence interval = 1.04-1.17); cardiovascular HR = 1.11 (1.01-1.22); ischaemic heart disease HR = 1.13 (1.02-1.25); respiratory HR = 1.26 (1.02-1.28)]. The associations were largely unaffected by additional adjustment for area-level deprivation category. A less consistent and generally implausible pattern of cause-specific BS-mortality associations was found for inverse-distance averaging of BS concentrations at nearby monitoring sites. BS-mortality associations in the Collaborative cohort were weaker and not statistically significant.

Conclusions: The association between mortality and long-term exposure to BS observed in the Renfrew/Paisley cohort is consistent with hypotheses of how air pollution may affect human health. The dissimilarity in pollution-mortality associations for different exposure models highlights the critical importance of reliable estimation of exposures on intra-urban spatial scales to avoid potential misclassification bias.

What this paper adds:

What is known:

- Epidemiological evidence suggests that long-term exposure to combustion-related air pollutants has adverse effects on health (which are more substantial than the effects associated with short-term exposures) but is limited by scientific uncertainties concerning exposure misclassification and potential confounding.
- There are scientific and policy requirements for cohort studies assessing air pollution health impacts in the UK to assess the appropriateness of extrapolation of findings from studies in the USA and European countries.

What this study adds:

- Associations between mortality and long-term exposure to black smoke air pollution observed in this study add to, and are generally consistent with, the limited observational evidence available to examine hypotheses of the extent to which air pollution may affect human health.
- Dissimilarity in pollution-mortality associations for different exposure models highlights the critical importance of reliable estimation of exposures on intra-urban spatial scales to avoid potential misclassification bias.
- The extent of dissimilarities noted between exposure models re-emphasises that inadequate human exposure classification will continue to be the one of the most challenging issues to address in future environmental epidemiology research; which emphasises the value of development of sufficiently extensive intra-urban pollution monitoring datasets to support improved epidemiological assessment.

Background:

North American [1-4] and European epidemiological cohort studies [5-7] provide evidence that long-term exposure to air pollutants have adverse effects on health, which are more substantial than the effects associated with short-term exposures [8-10]. Earlier analyses relied on inter-urban variations in air pollution with discrete urban areas represented by single monitoring sites. Subsequent attention has focused on exposure assignment determined from intra-urban variations in pollutant concentrations [2]. Extrapolating the findings from cohort studies in the USA and European countries to the population of the UK may not be appropriate because of variations in population demographics, cultural factors, and pollutant mixes. Correspondingly there are clear scientific and policy requirements for similar investigations within the UK [8].

Scotland has one of the highest mortality rates in the world for coronary heart disease [11] and lung cancer [12]. These mortality rates have been intensively studied and related to well-characterised socio-economic, lifestyle and medical risk factors. The purpose of the work described here was to apply different exposure estimation methods to investigate if long-term exposure to air pollution contributes to excess cause-specific mortality, after adjusting for individual-level risk factors in two large cohorts in Scotland with mortality follow-up periods of up to 25 years.

Methods:

Cohort participants: The study used two of the Midspan prospective cohorts [13]. The Renfrew/Paisley cohort was recruited from residents aged 45 to 64 of two towns in West Central Scotland and comprised 78% of the target population with 15,402 participants recruited between 1972 and 1976 [14]. The Collaborative cohort comprised 7,028 participants from 27 workplaces

in central Scotland, recruited between 1970 and 1973 [15]. Incomplete and incorrect postcodes, which could not be converted to grid references, restricted the number of participants selected for the present study to 15,331 and 6,680 for Renfrew/Paisley and Collaborative cohorts respectively.

Participants in both cohorts completed a health-related questionnaire and attended a screening medical examination. The questionnaire collected data on gender, date of birth, marital status, smoking status, occupation and address (from which full postcode of residence and area-level deprivation category (DEPCAT) were derived) (Table 1). DEPCAT is evaluated at the level of postcode sectors (average population: 5,000) and is calculated from census statistics on proportion of population in households without access to a car, in overcrowded households, with the head of household in social class IV or V, and in households with unemployed men [16]. Social class was derived using the UK Registrar General's classification based on occupation at time of screening [17]. Body mass index was calculated from measurements of height and weight.. Blood pressure was measured and a blood sample collected for measurement of plasma cholesterol [14].

Diagnoses for causes of death were based on the International Classification of Diseases (9^{th} revision). Five outcome mortality classifications were used: all cause (all codes), cardiovascular (410-414, 426-429, 434-440, 786.5), ischaemic heart disease (410-414), respiratory (480-487, 490-496, 786.0, 786.2) and lung cancer (162). These cause of death groupings were chosen to be compatible with related studies of effects of short-term pollution exposure [18]. Follow up for date and cause of death was maintained until the end of 1998.

Exposure modelling and assignment: Participants in the Collaborative cohort were geographically dispersed throughout cities, towns and villages in the central part of Scotland (Figure 1). The Renfrew/Paisley cohort participants were resident in a more localised area on the west side of the Glasgow conurbation. To provide an indication of geographical scale the contiguous conurbation of Glasgow, Paisley and Renfrew can be encompassed within a radius of 12 km, with Renfrew and Paisley encompassed by radii of 1.5 and 3.5 km respectively within this 12 km radius.

Daily black smoke (BS) measurements at 181 monitoring sites were obtained from the UK National Air Quality Archive [19]. BS is a metric of the optical darkness of airborne particulate matter collected on filter media [20]. Although quantified in units of $\mu g\ m^{-3}$ BS concentrations do not equate directly to the mass of any particular size fraction of airborne particulate matter. However, consistent standard calibrations (e.g. DETR [21]) have been used for many decades to convert reflectance to nominal concentration such that BS data are important measures of historic levels of air pollution used widely in epidemiological studies. The DETR [21] calibration procedures were used in the computation of UK government archived BS data used in the present manuscript. The use of the BS metric as a measure of particulate matter air pollution is well-established in the epidemiological research community and has been shown to be a good marker for traffic and other primary combustion-related urban air pollution often at least as predictive of negative health outcomes as PM_{10} or $PM_{2.5}$ [22].

These BS data were collected at a time when there was a move away from using coal as the main source of domestic heating fuel under the implementation of the UK Clean Air Acts. Quantitative emissions data for this period are relatively limited in detail. The location and times

of operation of the sites were at the discretion of the local authorities and the central government agency responsible for air pollution. A substantial amount of BS data was missing at several sites. Three approaches were used to estimate average long-term exposure to BS between 1970 and 1979 for individual cohort participants. There were insufficient pollutant observations to model exposure in the 1980s and 1990s at the same spatial resolution [23].

In the first approach, local knowledge of the geography and meteorological conditions in Scotland was used to allocate monitoring sites to 15 geographic regions (Figure 1). Each region had ≥ 1 site with ≥ 60% available BS data. The following model was used to impute missing data and compute geometric mean daily BS for 1970-79 at sites within each region,

$$y_{ij} = \ln(BS_{ij} + 0.5) = s_i + \beta_1 t_{ij} + \beta_2 t_{ij} I(t_{ij} > t^*) + day(t_{ij}) + month(t_{ij}) + \varepsilon_{ij} \quad (1)$$

Where i indexes the sites and j indexes the observations within a site, t_{ij} is time measured in days from 1/1/70; t^* is time from 1/1/75; $I(t_i > t^*) = 1$ when $t_i > t^*$ and 0 otherwise; s_i is a site specific intercept; $day(t_{ij})$ and $month(t_{ij})$ are factors for day of week and month respectively; and ε_{ij} is an error term. Geometric mean daily BS exposure (1970-79) was estimated for each cohort individual using an inverse distance weighted average of geometric mean BS at the nearest (< 1 km) monitoring sites. If there were no sites within 1 km exposure was assigned the weighted average of the two nearest monitoring sites. In this method, cohort individuals' assigned exposure could only range from the minimum to the maximum of the nearest sites.

In the second approach, the 10-year average BS at each site (after imputation of missing values as before) was related to four local air quality predictors (LAQP): altitude above sea level (*A*); household density (*HD*) within a 250m buffer [24]; distance to nearest major road (*MR*) (motorways and 'A' roads in 2001 from the National Atmospheric Emissions Inventory [25]);

and distance to the edge of the nearest urban boundary (*UB*) (derived from Ordnance Survey data). The model also included an indicator (*UA.Ind*) of whether the monitoring site was inside or outside a small (<17.7 km^2; cut-point defined by median area of urban areas containing monitoring sites) or large urban area. Five spatial regression models were examined in sensitivity analyses [23]. The most parsimonious configuration was a semi-parametric model with bivariate smooth trend of geographical coordinates, *s*(*E,N*), and parametric terms for LAQP, which was then used to predict 10-year average BS at each residential location.

$$\ln(BS+0.5) = s(E, N) + \beta_1 A + \beta_2\sqrt{HD} + \beta_3\sqrt{MR} + \beta_4 UB + \beta_5 UA.Ind \qquad (2)$$

where $\beta_1 \ldots \beta_5$ are fixed effects parameters for LAQP.

The most detailed approach, multilevel spatio-temporal modelling (MultiBS), employed a combination of time series, imputation and spatial smoothing techniques to model the change in monthly BS simultaneously taking into account seasonal effects and LAQP.

$$y_{ij} = f(t_{ij}) + g_i(t_{ij}) + \alpha_{1c} Cos\left(\frac{t_{ij}}{12}\right) + \alpha_{1s} Sin\left(\frac{t_{ij}}{12}\right) + \alpha_2 A_i + \alpha_3\sqrt{HD_i} + \alpha_4\sqrt{MR_i} + \alpha_5 UB_i + \varepsilon_{ij} \qquad (3)$$

Here *i* indexes the sites and *j* indexes the temporal observations; t_{ij} is the number of months from January 1970; $f(t_{ij})$ is the BS temporal trend averaged over the population of all sites; $g_i(t_{ij})$ is the deviation of the i^{th} site from the population mean at time t_{ij} ($f(t_{ij})$ and $g_i(t_{ij})$ were modeled flexibly using penalised linear splines); sine and cosine terms model monthly seasonal effects with α_{1c}, α_{1s} as fixed-effect parameters; and $\alpha_2 \ldots \alpha_5$ are fixed-effect parameters of LAQP.

The AMBS and MultiBS models were similar in that both used spatial smoothing to estimate participants' exposure by taking into account both air pollution concentrations at monitoring sites nearby their residences and local environmental determinants by means of LAQP. However the

multilevel model has the ability to estimate coefficients between BS and LAQP in the presence of missing data, and hence was not dependent on the imputation techniques used to replace missing data in the first two approaches. BS exposures calculated by the three techniques are subsequently referred to as IDWBS, AMBS and MultiBS. Further details of the development, evaluation and application of these exposure models are given in Robertson et al. [23].

We included only participants who lived within 5 km of the nearest sites for all three exposure models. Estimated 1970-79 geometric mean exposure concentrations at participants' residential addresses in Renfrew/Paisley ranged from 14.9 to 27.1 μg m^{-3}, 5.9 to 24.4 μg m^{-3}, and 6.4 to 28.7 μg m^{-3} for IDWBS, AMBS and MultiBS respectively; and in Collaborative ranged from 5.4 to 70.0 μg m^{-3}, 6.2 to 48.5 μg m^{-3}, and 4.6 to 55.3 μg m^{-3} respectively (Table 1).

The three exposure models were evaluated in a cross-validation study [23]. Monitoring sites with > 80% data coverage were identified and any missing data was imputed with a site-specific time-series model with a flexible trend, month and day effects to give 39 sites with 'complete' data. Ten-year mean BS concentrations at these 39 sites ranged from 8.9 μg m^{-3} to 48.2 μg m^{-3}. We then created a 'test data set' from 19 of these sites, selected at random, and a 'training data set' from the remaining 20 sites together with all the sites with < 80% data coverage. The model was fitted to the training data set (of 162 sites) and then used to predict BS in the test data set. This cross-validation procedure was repeated 10 times with different random selections from the 39 complete data sites forming the test set. The average mean squared differences on the log black smoke scale [(ln μg m^{-3})^2] were 0.171, 0.171, and 0.090 for IDWBS, AMBS and MultiBS exposure models respectively.

Interpolated maps of BS concentrations were prepared from estimated BS exposure at address postcodes of cohort participants [23]. The AMBS and MultiBS models provided a much more consistent and (from local knowledge) plausible prediction of exposure at addresses of individuals than the IDWBS model (e.g. the IDWBS model failed to predict anticipated lower concentrations for many cohort addresses in residential areas in south Paisley as all estimates were constrained to remain within the high concentrations measured in the centre of Paisley). Collectively the cross-validation and map visualisation suggest that the LAQP-based models produce a more realistic prediction of likely exposures in the cohorts [23].

Survival Analysis: Associations between estimated long-term exposure to air pollution using the three different exposure models and cause-specific mortality were examined using Cox proportional hazards regression, with baseline hazard functions stratified by 1 year age groups and gender, for follow-up to the end of 1998.

Baseline variables included: marital status; smoking status (never, ex-smoker, or current smokers who smoked 1-14, 15-24, 25+ cigarettes per day, pipe or cigar smokers); social class (categorised as I:high, II, III non-manual, III manual, IV, V:low); body mass index (expressed in quintiles); systolic blood pressure; and cholesterol. The latter two variables were used only for modeling all cause, cardiovascular and ischaemic heart disease mortality.

Participants with missing systolic blood pressure and cholesterol were removed. 423 (2.8%) participants with missing social class in Renfrew/Paisley were recoded as a separate level in social class. As there were only 11 (0.2%) participants with missing social class in Collaborative, they were removed from the analyses.

To investigate sensitivity to possible additional confounding by area-level socio-economic status, additional adjustments for DEPTCAT (1 (least deprived), 2, …., 7 (most deprived)) were included. Further sensitivity analyses used shared gamma frailty models, where a random effect cluster was applied to postcode sectors and deprivation categories. Additional sensitivity analyses examined the effect of exclusion of participants who lived at different distances from monitoring sites (for IDWBS) and different follow-up time periods (all exposure models). Possible BS effect modification by gender, smoking status, body mass index and social class was also examined.

Statistical analyses were performed using SPLUS 7.0, R 2.14.1 and SPSS 12.0.1. A R package "frailtypack" was used to fit the shared frailty model with parameters estimated by penalised likelihood maximisation.

Results:

In analyses of the Renfrew/Paisley cohort (Table 2(A)) the adjusted hazard ratios for all-cause and cause-specific mortality, attributable to an increment of 10 μg m^{-3} long-term MultiBS exposure in 1970-79, were, in descending order: respiratory (hazard ratio, HR = 1.26, 95% CI = 1.02-1.55); ischaemic heart disease (HR = 1.13, 1.02-1.25); cardiovascular (HR = 1.11, 1.01-1.22); all cause (HR = 1.10, 1.04-1.17) and lung cancer (HR = 1.00, 0.84-1.20) for follow-up till end of 1998. Associations between BS and all-cause and specific causes of death were slightly attenuated but persisted with additional adjustment of deprivation category in the standard Cox model.

Effect magnitudes for exposures predicted via AMBS were similar to those via MultiBS. In contrast, associations between IDWBS estimates and mortality were markedly different from the AMBS and MultiBS models, with inconsistent directions and relatively large confidence intervals.

In the Collaborative cohort BS-mortality associations were lower and not significantly elevated, except for lung cancer [HR = 1.11 (0.96-1.30) for MultiBS]. and were similar for different exposure models (Table 2(B)).

The hazard ratios in both cohorts were largely unaffected by stepwise adjustment of risk factors (Table 3).

In sensitivity analyses using shared gamma frailty models there was no evidence for the Collaborative cohort of heterogeneity between postcode sectors or between deprivation categories. In Renfrew/Paisley, a much smaller geographic area with only 14 postcode sectors, there was evidence for heterogeneity between postcode sectors (deduced from a modified Wald test: variance of the random effect of 0.00637 divided by standard error 0.0335, which gave 1.82 > 1.64, critical value of one-sided normal test) but no evidence of heterogeneity between deprivation categories. Inclusion of a shared frailty component associated with deprivation category was not required as the risk factors in the model accounted for most of the variance in survival time.

There was no evidence of BS effect modification by gender, smoking or social class for all cause and specific causes of death for both cohorts (results not shown).

Discussion:

BS-mortality associations were observed in the geographically localised Renfrew/Paisley cohort for all-cause, cardiovascular, ischaemic heart disease and respiratory mortality. Associations between BS and cause-specific outcomes were generally consistent for the LAQP-based MultiBS and AMBS exposure models, with a less consistent and generally implausible pattern of associations noted for the IDWBS exposure model. There was limited evidence of possible pollution-related effects in separate analyses of the Collaborative cohort.

Analyses based on the MultiBS exposure model in Renfrew/Paisley indicated highest hazard ratios for respiratory, followed by ischaemic heart disease, cardiovascular and all-cause mortality. BS-mortality associations for the AMBS exposure model were similar to those for MultiBS. These findings are consistent with hypotheses of how air pollution may affect human health [10 26] and the limited evidence base (reviewed by [22]) on BS-mortality associations from cohort [5 7] and cross-sectional studies [27] (Table 4). For example, observation of all-cause mortality HR of 1.10 (95% CI: 1.04-1.17) associated with an increment of 10 $\mu g\ m^{-3}$ long-term MultiBS in this study is consistent with equivalent effect magnitudes in 2 other cohort studies that use BS as an exposure metric: in the Netherlands (all-cause mortality HR of 1.05 (1.00-1.11) [7]); and in France (all-cause mortality HR of 1.07 (1.03-1.10) [5]). These cohort-based risk estimates for all-cause mortality appear to be higher than similar risk estimates made in a small area ecological study in the UK (e.g. all-cause mortality HR = 1.019 (1.018-1.021) and 1.007 (1.006-1.009) for analyses before and after adjustment for area-level deprivation for 0-8 year exposure window [27]). Relatively large BS-respiratory mortality associations are evident in the Midspan, NLCS-

Air and GB small area studies (respiratory mortality was not examined separately in the PAARC study).

The IDWBS was anticipated *a priori* to be inadequate for estimation of the effects of local road traffic and household emissions (particularly domestic coal fires before more extensive implementation of smoke control areas under the UK Clean Air Acts) and dispersion and advection processes. This may explain why IDWBS exposure estimates resulted in unexpected 'protective' BS-mortality associations in Renfrew/Paisley, although these were less evident when analysis was restricted to cohort participants within 2 km of a monitoring site (results not shown). This suggests that IDWBS estimation results in gross exposure misclassification. For example, local knowledge of pollution climates suggests that the IDWBS model substantially overestimated actual exposures in suburban areas in south Paisley by assigning cohort participants with a distance weighted average of the means of the two nearest monitoring sites in relatively polluted parts of the centre of Paisley. In contrast the use of LAQP provided consistent and plausible estimates of intra-urban variations in BS exposure as predicted concentrations were not constrained to lie within the range of concentrations observed at the nearest (but not necessarily 'near') monitoring sites [23].

In the reanalysis of the ACS and Harvard Six Cities cohorts, which assigned exposure based on community average concentrations, the hazard ratio for respiratory mortality was lower in magnitude than that for cardiovascular mortality with relatively wide confidence intervals [28]. In contrast, in this study of the Renfrew/Paisley cohort using LAQP-based exposure estimation, the hazard ratio for respiratory mortality was higher in magnitude than for any of the other outcomes.

There are limitations in this study of the Midspan cohorts that are shared to a greater or lesser extent with most, if not all, cohort studies of long-term exposure to air pollution. These include exposure misclassification (through missing exposure data, limitations of the exposure model in capturing long-term personal exposures of multiple pollutant metrics that may be relevant to the outcomes being studied, and lack of information on participants' mobility) and potentially incomplete adjustment for confounding (through unknown individual and area-level risk factors).

The reasons for the weaker BS-mortality associations in the Collaborative compared to Renfrew/Paisley cohort remain speculative. The MultiBS estimated BS concentrations in the Renfrew/Paisley cohort area had a relatively small IQR and range of 6 and 6–29 μg m^{-3} respectively (compared to 8 and 5-55 μg m^{-3} for the Collaborative cohort), but the effect magnitudes for Renfrew/Paisley were more elevated for all specific causes except lung cancer. Exposure misclassification, lower number of participants/events, lower susceptibility, and/or unmeasured confounding factors may have been important. It is possible that the non-occupational nature of the Renfrew/Paisley cohort increased the number of individuals who spent more time at their residential address reducing exposure misclassification.

Although the risks of air pollution on lung cancer in the Collaborative cohort were estimated relatively imprecisely because of low number of events, the direction and magnitude of the effect estimated from the MultiBS exposure model (HR = 1.11, CI = 0.97-1.30 for follow up to 1998) were not inconsistent with the 8% increase risk in lung cancer mortality noted in extended analyses of the ACS cohort [29]. It is possible that high smoking rates in the Renfrew/Paisley cohort (>80% of males with history of smoking and the relatively high overall consumption of

cigarettes) may have obscured associations between lung cancer mortality and air pollution estimated from the AMBS and MultiBS exposure models. For reasons that remain unclear, relatively high rates of lung cancer were noted in suburban areas in the south west of the Renfrew/Paisley study area compared to lower rates of lung cancer in the town centre areas of Paisley and Renfrew. Cohort participants in south west Paisley would have been assigned a distance weighted average of the means of the two nearest monitoring sites in relatively polluted parts of the centre of Paisley. As noted above, the IDWBS exposure estimates are prone to this type of limitation compared to the AMBS and MultiBS exposure models and correspondingly the associations observed between lung cancer mortality and IDWBS in the Renfrew/Paisley cohort (Table 2) were considered to be anomalous.

Participants' exposures were based on their residential addresses recorded at recruitment in the 1970s. Information on relocations and recent addresses was unavailable but linkage of the Renfrew/Paisley cohort to a national patient database for Scotland suggested that the majority of survivors (84%) were still resident in the Argyll and Clyde Health Board area in the West of Scotland in 1995 [30]. This does not however provide information on the extent to which participants may have changed address within this Health Board area and/or within the urban conurbation of Renfrew/Paisley. Thus exposure misclassification may have resulted from a lack of information about participants' mobility.

Information regarding some potential risk factors, including smoking and body mass index, were only obtained at the time of recruitment. Hence, adjustments for changes in these factors, which might alter the risks of air pollution on health, could not be made. Information on education level was available for the Collaborative cohort, but not for the Renfrew/Paisley cohort. However,

Davey Smith et al. [15] have shown that occupational social class can be a stronger predictor of health outcomes than education. Additionally, there is evidence that underlying social inequalities in health in the UK may be related more clearly to current social circumstances rather than childhood circumstances [15].

The LAQP exposure models predicting air pollution concentrations include household density and distance to nearest road variables that may be associated with mortality independently of their association with air pollution (as indirect measures of area-level socio-economic conditions). This raises the possibility of confounding [31]. However, as pollution climate is highly (and highly plausibly) dependent on the LAQP variables, inclusion of these variables in the survival model presents substantial risk of 'over-adjustment' and possible obscuration of genuine pollution effects. Given that direct individual measures of long-term pollution exposure are unfeasible (and impossible retrospectively), the estimation of long-term effects of air pollution requires a choice between: (a) definite, and possibly gross, exposure misclassification due to very poor estimation of individual exposures; or (b) more precise estimation of individual exposures that may entail possible confounding by local air quality predictors. The potential for unknown confounding mechanisms accounting for the apparent effect of long-term exposure to particles on mortality is likewise recognised in reviews of scientific issues in air pollution and health research [10].

To test further for possible socio-economic confounding an additional area-level measure of deprivation was added to the standard Cox model [32]. This measure has been shown to be related to multiple heath outcomes in the Renfrew/Paisley cohort [33] and to modify observed effects of short-term exposure to BS in the general population from which the cohort was

sampled [34]. Additional adjustment for DEPCAT slightly attenuated the associations between long-term BS exposure and mortality (Table 2 Panels C & D & Table 3), but the overall pattern of association remained broadly consistent with the analyses prior to adjustment for DEPCAT and with the magnitude of pollution effects published in a UK review of the health effects of long-term exposure to air pollution (best estimates [and 95% CI] of relative risk of: 1.06 [1.02-1.11]; 1.09 [1.03-1.16]; 1.08 [1.01-1.16] for all-cause; cardiopulmonary; and lung cancer mortality respectively [8]). Further sensitivity analyses using a shared frailty model revealed similar pollution-mortality associations as in the standard Cox models, with no evidence of confounding by DEPCAT.

The interpretation of this additional adjustment is similarly complicated by the possibility of 'over-adjustment' that may obscure underlying effects of pollution exposure as the area level DEPCAT variable is partly derived from individual level social class variables, and by the possibility that air pollution has a role in the contextual effect of neighbourhood-level deprivation on mortality [35-37]. Complexities of this nature may have contributed to inconsistent evidence found in reviews of the effect of socioeconomic status on the relationship between air pollution exposure and health [38-39] and are subject to ongoing research developments (e.g. using multilevel analytical approaches) which continue to face conceptual and methodological challenges to establishing causal inference [40]. The detailed pollution exposure estimates produced for the Midspan cohorts may provide a useful dataset for future research as methodological progress is made. In the meantime it is re-emphasised that retrospective individual-level exposure estimation is just the best possible estimate of individual exposure within constraints of currently available information and analytical approaches and that the data in panels A and C, and B and D of Table 2 provides a range of HR estimates of the effects of air

pollution between possible under- and over-adjustment for confounding in the combined exposure and survival models.

Conclusions:

The associations between mortality and long-term exposure to BS observed in this study in the UK are broadly consistent with previous evidence from other countries and hypotheses of how particulate matter air pollution may affect human health. The dissimilarity in health effects based on different exposure models highlights the critical importance of reliable estimation of long-term exposures on a fine intra-urban spatial scale to avoid potential misclassification problems inherent in air pollution epidemiology studies. The extent of the dissimilarities noted between exposure models re-emphasises that it is likely that inadequate human exposure classification will continue to be the one of the most challenging issues to address in future environmental epidemiology research; and this has important implications for the development of fit-for-purpose pollution monitoring and modelling capabilities by local and central government and their health protection agencies.

Acknowledgement: The authors of this paper gratefully acknowledge funding from Department of Health (England) Policy Research Programme as part of the Initiative on Air Pollution (Research Grant 0020015). The authors are also particularly grateful for the encouragement and support given by the late David Hole former Professor of Epidemiology and Biostatistics & Head of West of Scotland Cancer Surveillance Unit in his enthusiastic role in planning and implementing this research.
Disclaimer: The views in this paper are those of the authors but not necessarily the Department of Health (England).
Competing interest statement: No competing interests.

References:

1 Laden F, Schwartz J, Speizer FE et al. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am.J.Respir.Crit Care Med.* 2006;173:667-672.

2 Krewski D, Jerrett M, Burnett RT et al. Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality. HEI Research Report 140. Health Effects Institute, Boston, MA. http://pubs.healtheffects.org/view.php?id=315 Accessed 2012. 2009.

3 Chen LH, Knutsen SF, Shavlik D et al. The Association between Fatal Coronary Heart Disease and Ambient Particulate Air Pollution: Are Females at Greater Risk? *Environ Health Perspect* 2005;113.

4 Puett RC, Hart JE, Yanosky JD et al. Chronic Fine and Coarse Particulate Exposure, Mortality, and Coronary Heart Disease in the Nurses' Health Study. *Environmental Health Perspectives* 2009;117:1697-1701.

5 Filleul L, Rondeau V, Vandentorren S et al. Twenty five year mortality and air pollution: results from the French PAARC survey. *Occup Environ Med* 2005;62:453-460.

6 Nafstad P, Haheim LL, Wisloff T et al. Urban air pollution and mortality in a cohort of Norwegian men. *Environmental Health Perspectives* 2004;112:610-615.

7 Beelen R, Hoek G, van den Brandt PA et al. Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environmental Health Perspectives* 2008;116:196-202.

8 COMEAP. Long-Term Exposure to Air Pollution: Effect on Mortality. Final Report by: Committee on the Medical Effects of Air Pollutants (COMEAP), June 2009. http://www.dh.gov.uk/ab/COMEAP/DH_108151 2009.

9 Pope CA. Mortality effects of longer term exposures to fine particulate air pollution: Review of recent epidemiological evidence. *Inhalation Toxicology* 2007;19:33-38.

10 Pope CA, Dockery DW. Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association* 2006;56:709-742.

11 Scottish-Executive-Health-Department. Coronary Heart Disease / Stroke Task Force Report: Scottish Executive 2001.

12 Parkin DM, Whelan SL, Ferlay J et al. *Cancer Incidence in Five Continents Vol VII* International Agency for Research on Cancer, 1997.

13 Hart CL, MacKinnon PL, Watt GC et al. The midspan studies. *International Journal of Epidemiology* 2005;34:28-34.

14 Hawthorne VM, Watt GCM, Hart CL et al. Cardiorespiratory disease in men and women in urban Scotland: baseline characteristics of the Renfrew/Paisley (Midspan) study population. *Scott Med J* 1995;40:102-107.

15 Davey Smith G, Hart CL, Hole DJ et al. Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Community Health* 1998;52:153-160.

16 Carstairs V. Deprivation indices: Their interpretation and use in relation to health. *Journal of Epidemiology and Community Health* 1995;49:S3-S8.

17 Hart CL, Hole DJ, Gillis CR et al. Social class differences in lung cancer mortality: risk factor explanations using two Scottish cohort studies. *Int J Epidemiol* 2001;30:268-274.

18 Beverland IJ, Cohen GR, Carder M et al. A novel comparison of short-term and long-term air pollution exposure associations with mortality in two cohorts in Scotland. *Env Health Perspect.;* Accepted for publication.

19 DEFRA. National air quality data archive. Department for Environment Food & Rural Affairs (DEFRA) http://www.airquality.co.uk/ [accessed 1 October 2005]. 2005.

20 Heal MR, Quincey P. The relationship between black carbon concentration and black smoke: A more general approach. *Atmospheric Environment* 2012;54:538-544.

21 DETR. Department of Environment Transport & Regions (DETR) Instruction Manual: UK Smoke and Sulphur Dioxide Network, AEAT-1806. AEA Technology, Harwell. 1999.

22 Janssen NAH, Hoek G, Simic-Lawson M et al. Black Carbon as an Additional Indicator of the Adverse Health Effects of Airborne Particles Compared with PM(10) and PM(2.5). *Environmental Health Perspectives* 2011;119:1691-1699.
23 Beverland IJ, Robertson C, Yap C et al. Comparison of Models for Estimation of Long-Term Exposure to Air Pollution in Two Cohorts in Scotland. *Atmospheric Environment Submitted*.
24 SURPOP. Modelled household count in 200 m grid squares from the 1981 census. Source: The 1991 Census, Crown Copyright. ESRC/JISC purchase. http://www.census.ac.uk/cdu/surpop/ 2006.
25 NAEI. National Atmospheric Emissions Inventory (NAEI): Annual average daily traffic flows at count points on major roads (Scotland). http://www.naei.org.uk/ 2005.
26 Brunekreef B, Holgate ST. Air pollution and health. *The Lancet* 2002;360:1233-1242.
27 Elliott P, Shaddick G, Wakefield JC et al. Long-term associations of outdoor air pollution with mortality in Great Britain. *Thorax* 2007;62:1088-1094.
28 Krewski D, Burnett R, Jerrett M et al. Mortality and long-term exposure to ambient air pollution: ongoing analyses based on the American Cancer Society cohort. *J.Toxicol.Environ.Health A* 2005;68:1093-1109.
29 Pope CA, Burnett RT, Thun MJ et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama-Journal of the American Medical Association* 2002;287:1132-1141.
30 Hanlon P, Walsh D, Whyte BW et al. The link between major risk factors and important categories of admission in an ageing cohort. *Journal of Public Health Medicine* 2000;22:81-89.
31 Moore DK, Jerrett M, Mack WJ et al. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *Journal of Environmental Monitoring* 2007;9:246-252.
32 Hoek G, Beelen R, de Hoogh K et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* 2008;42:7561-7578.
33 Davey Smith G, Hart C, Watt G et al. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley study. *Journal of Epidemiology and Community Health* 1998;52:399-405.
34 Carder M, McNamee R, Beverland I et al. Does deprivation index modify the acute effect of black smoke on cardiorespiratory mortality? *Occupational and Environmental Medicine* 2010;67:104-110.
35 Pearce JR, Richardson EA, Mitchell RJ et al. Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. *Transactions of the Institute of British Geographers* 2010;35:522-539.
36 Naess O, Piro FN, Nafstad P et al. Air Pollution, Social Deprivation, and Mortality A Multilevel Cohort Study. *Epidemiology* 2007;18:686-694.
37 Evans GW, Kantrowitz E. Socioeconomic status and health: The potential role of environmental risk exposure. *Annual Review of Public Health* 2002;23:303-331.
38 Deguen S, Zmirou-Navier D. Social inequalities resulting from health risks related to ambient air quality-A European review. *European Journal of Public Health* 2010;20:27-35.
39 Laurent O, Bard D, Filleul L et al. Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. *Journal of Epidemiology and Community Health* 2007;61:665-675.
40 Kawachi I, Subramanian SV. Neighbourhood influences on health. *Journal of Epidemiology and Community Health* 2007;61:3-4.

Table 1 Summary statistics of individual baseline mortality data in Renfrew/Paisley and Collaborative cohorts (sd: standard deviation; LQ: lower quartile; UQ: upper quartile; min: minimum, max: maximum)

Variables		Renfrew/Paisley (*n* = 15,331)	Collaborative (*n* = 6,680)
Age in years	mean (sd)	54.3 (5.6)	48.1 (6.6)
	median [LQ, UQ]	54.0 [50.0,59.0]	48.0 [43.0,53.0]
	(min, max)	(45.0,64.0)	(35.0,64.0)
Gender	female (%)	8324 (54.3%)	987 (14.8%)
	male (%)	7007 (45.7%)	5693 (85.2%)
Marital Status	married (%)	12284 (80.1%)	5852 (87.6%)
	other (%)	271 (1.8%)	60 (0.9%)
	single (%)	1411 (9.2%)	582 (8.7%)
	widowed (%)	1365 (8.9%)	185 (2.8%)
	missing (%)	0 (0.0%)	1 (0.0%)
Social class	I (high) (%)	538 (3.5%)	767 (11.5%)
	II (%)	2221 (14.5%)	1104 (16.5%)
	IIIN (%)	2798 (18.3%)	1813 (27.1%)
	IIIM (%)	4289 (28.0%)	1216 (18.2%)
	IV (%)	3758 (24.5%)	1539 (23.0%)
	V (low) (%)	1304 (8.5%)	230 (3.4%)
	missing (%)	423 (2.8%)	11 (0.2%)
Smoking habits	never (%)	4983 (32.5%)	1343 (20.1%)
	ex (%)	2343 (15.3%)	1440 (21.6%)
	1-14 (%)	2394 (15.6%)	981 (14.7%)
	15-24 (%)	4057 (26.5%)	1891 (28.3%)
	25+ (%)	1407 (9.2%)	868 (13.0%)
	Pipe-cigar (%)	147 (1.0%)	154 (2.3%)
	missing (%)	0 (0.0%)	3 (0.0%)
Body mass index, bmi (kg/m^2)	mean (sd)	25.8 (4.0)	25.1 (3.2)
	median (LQ, UQ)	25.5 [23.1,28.0]	25.0 [23.0,27.0]
	(min, max)	(14.1,67.9)	(15.7,51.7)
	missing (%)	14 (0.1%)	1 (0.0%)
Systolic blood pressure (mm Hg)	mean (sd)	149.4 (24.4)	134.4 (18.2)
	median (LQ, UQ)	146.0 [132.0,164.0]	131.0 [122.0,143.0]
	(min, max)	(80.0,270.0)	(90.0,257.0)
	missing (%)	7 (0.0%)	5 (0.1%)
Cholesterol (mmol/l)	mean (sd)	6.2 (1.1)	5.9 (1.0)
	median (LQ, UQ)	6.1 [5.4,6.8]	5.8 [5.2,6.5]
	(min, max)	(1.9,20.5)	(1.9,11.5)
	missing (%)	137 (0.9%)	41 (0.6%)
Deprivation Category	1 (least deprived)	967 (6.3%)	616 (9.2%)
	2	0 (0.0%)	536 (8.0%)
	3	2086 (13.6%)	966 (14.5%)
	4	3349 (21.8%)	1056 (15.8%)
	5	5539 (36.1%)	1425 (21.3%)
	6	2765 (18.0%)	1370 (20.5%)
	7 (most deprived)	625 (4.1%)	709 (10.6%)
	missing(%)	0 (0.0%)	2 (0.0%)
Specific group causes of death up to end of 1998	Dead (%)	7767 (50.7%)	2885 (43.2%)
	Cardiovascular (%)	3041 (19.8%)	1169 (17.5%)
	Ischaemic Heart Disease (%)	2534 (16.5%)	954 (14.3%)
	Respiratory (%)	606 (4.0%)	178 (2.7%)
	Lung Cancer (%)	798 (5.2%)	288 (4.3%)
Distance from nearest BS monitor (km)	mean (sd)	1.7 (1.1)	1.8 (2.3)
	median (LQ, UQ)	1.4 [0.8,2.6]	1.2 [0.7,2.0]
	(min, max)	(0.0,4.6)	(0.0,69.5)
Predicted black smoke exposure using IDWBS ($\mu g\ m^{-3}$) (within 5 km of monitoring sites)	mean (sd)	23.6 (2.8)	22.6 (8.3)
	median [LQ,UQ]	25.2 [21.6,25.7]	21.3 [16.7,26.8]
	(min,max)	(14.9,27.1)	(5.4,70.0)
	missing (%)	0 (0%)	449 (6.7%)
Predicted black smoke exposure using AMBS ($\mu g\ m^{-3}$) (within 5 km of monitoring sites)	mean (sd)	18.8 (2.7)	23.0 (7.6)
	median [LQ,UQ]	19.4 [16.8,20.9]	21.5 [17.6,27.8]
	(min,max)	(5.9,24.4)	(6.2,48.5)
	missing (%)	0 (0%)	422 (6.3%)
Predicted black smoke exposure obtained via MultiBS ($\mu g\ m^{-3}$) (within 5 km of monitoring sites)	mean (sd)	19.3 (3.9)	23.2 (7.5)
	median [LQ,UQ]	19.8 [16.1,22.4]	21.8 [18.3,26.6]
	(min,max)	(6.4,28.7)	(4.6,55.3)
	missing (%)	0 (0%)	364 (5.4%)

Table 2: Adjusted hazard ratios per 10 μg m^{-3} increment of geometric mean black smoke concentration for 1970-79, with corresponding 95% CI for all cause and cause-specific mortality for follow-up till end of 1998 for the Renfrew/Paisley (panel A) and Collaborative (panel B) cohorts. Hazard ratios in Panels (A) and (B) were estimated by Cox proportional hazards regression model adjusted for baseline risk factors (listed in Table 3). Panels (C) and (D) outline effect of adjustment for area-level Deprivation Category in addition to above baseline risk factors.

Exposure Model[a]	Cause[b]	Cases	N	HR	95% CI	p	Cases	N	HR	95% CI	p
				(A) Renfrew/Paisley					**(B) Collaborative**		
MultiBS	All Cause	7691	15188	1.10	(1.04-1.17)	0.002	2711	6257	1.01	(0.96-1.06)	0.75
	CVD	3014	15188	1.11	(1.01-1.22)	0.028	1091	6257	1.03	(0.95-1.12)	0.48
	IHD	2512	15188	1.13	(1.02-1.25)	0.019	890	6257	1.03	(0.94-1.12)	0.56
	Respiratory	606	15331	1.26	(1.02-1.55)	0.035	174	6299	0.97	(0.79-1.18)	0.76
	Lung Cancer	798	15331	1.00	(0.84-1.20)	0.97	273	6299	1.11	(0.96-1.30)	0.17
AMBS	All Cause	7691	15188	1.14	(1.04-1.24)	0.003	2687	6200	1.01	(0.96-1.06)	0.79
	CVD	3014	15188	1.14	(1.00-1.31)	0.052	1082	6200	1.02	(0.95-1.11)	0.55
	IHD	2512	15188	1.19	(1.02-1.37)	0.023	883	6200	1.01	(0.92-1.10)	0.83
	Respiratory	606	15331	1.43	(1.05-1.96)	0.023	173	6241	0.93	(0.76-1.14)	0.49
	Lung Cancer	798	15331	0.98	(0.76-1.26)	0.85	269	6241	1.15	(0.98-1.34)	0.082
IDWBS 5km	All Cause	7691	15188	0.90	(0.83-0.98)	0.014	2677	6174	1.00	(0.95-1.05)	0.98
	CVD	3014	15188	0.77	(0.67-0.87)	<0.001	1078	6174	1.02	(0.95-1.10)	0.61
	IHD	2512	15188	0.81	(0.70-0.93)	0.003	880	6174	1.01	(0.93-1.10)	0.80
	Respiratory	606	15331	0.94	(0.71-1.26)	0.70	172	6214	0.96	(0.80-1.17)	0.71
	Lung Cancer	798	15331	1.33	(1.03-1.74)	0.032	266	6214	1.07	(0.93-1.24)	0.34
				(C) Renfrew/Paisley					**(D) Collaborative**		
MultiBS	All Cause	7691	15188	1.08	(1.02-1.15)	0.015	2710	6255	1.01	(0.95-1.06)	0.82
	CVD	3014	15188	1.10	(1.00-1.22)	0.060	1091	6255	1.02	(0.93-1.10)	0.72
	IHD	2512	15188	1.12	(1.00-1.25)	0.050	890	6255	1.01	(0.92-1.11)	0.87
	Respiratory	606	15331	1.30	(1.04-1.63)	0.024	174	6297	0.96	(0.78-1.18)	0.69
	Lung Cancer	798	15331	0.99	(0.81-1.21)	0.92	273	6297	1.15	(0.98-1.34)	0.095
AMBS	All Cause	7691	15188	1.12	(1.02-1.23)	0.015	2686	6198	1.01	(0.96-1.06)	0.80
	CVD	3014	15188	1.14	(0.98-1.32)	0.089	1082	6198	1.02	(0.94-1.11)	0.67
	IHD	2512	15188	1.18	(1.00-1.39)	0.047	883	6198	1.00	(0.91-1.09)	0.93
	Respiratory	606	15331	1.57	(1.12-2.19)	0.009	173	6239	0.91	(0.74-1.12)	0.38
	Lung Cancer	798	15331	0.96	(0.72-1.28)	0.77	269	6239	1.18	(1.00-1.38)	0.047
IDWBS 5km	All Cause	7691	15188	0.86	(0.79-0.94)	0.001	2676	6172	1.00	(0.95-1.05)	0.91
	CVD	3014	15188	0.71	(0.61-0.82)	<0.001	1078	6172	1.02	(0.94-1.11)	0.59
	IHD	2512	15188	0.74	(0.63-0.87)	<0.001	880	6172	1.00	(0.92-1.09)	0.98
	Respiratory	606	15331	0.80	(0.57-1.11)	0.18	172	6212	0.93	(0.75-1.14)	0.47
	Lung Cancer	798	15331	1.34	(1.00-1.79)	0.050	266	6212	1.10	(0.95-1.28)	0.21

[a] MultiBS - Multilevel Black Smoke Model with local air quality predictors, with up to 294 and 141 per 1000 person years for Renfrew/Paisley and Collaborative respectively.
AMBS - Additive model using area based imputed data and local air quality predictors, with up to 294 and 140 per 1000 person years for Renfrew/Paisley and Collaborative respectively.
IDWBS 5km - Inverse distance weighted assignment using area based imputed data., with up to 294 and 139 per 1000 person years for Renfrew/Paisley and Collaborative respectively.
IDWBS 5km - Inverse distance weighted assignment, restricting to participants within 5 km, using area based imputed data, with up to 184 and 108 per 1000 person years for Renfrew/Paisley and Collaborative respectively.
[b] CVD - Cardiovascular Disease; IHD - Ischaemic Heart Disease.

Table 3. Adjusted hazard ratios per 10 μg m^{-3} increment of geometric mean black smoke concentration for 1970-79, with corresponding 95% CI for all cause and cause-specific mortality for the Renfrew/Paisley and Collaborative cohorts with stepwise adjustment of risk factors in Renfrew/Paisley and Collaborative with MultiBS.

	All Cause		CVD		IHD		Respiratory		Lung Cancer	
Model Covariates [a]	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Renfrew/Paisley										
black smoke only	1.12	(1.05-1.18)	1.09	(1.00-1.20)	1.12	(1.01-1.23)	1.34	(1.08-1.65)	1.07	(0.89-1.28)
+ marital status	1.1	(1.04-1.16)	1.08	(0.99-1.19)	1.1	(0.99-1.22)	1.31	(1.06-1.62)	1.07	(0.89-1.28)
+ body mass index	1.1	(1.03-1.16)	1.08	(0.99-1.19)	1.11	(1.00-1.23)	1.28	(1.03-1.58)	1.04	(0.87-1.24)
+ smoking	1.08	(1.02-1.15)	1.07	(0.98-1.18)	1.1	(0.99-1.22)	1.27	(1.02-1.56)	1.01	(0.85-1.21)
+ cholesterol	1.08	(1.02-1.15)	1.07	(0.97-1.17)	1.09	(0.98-1.21)	NA	NA	NA	NA
+ systolic blood pressure	1.11	(1.05-1.18)	1.12	(1.02-1.24)	1.14	(1.03-1.26)	NA	NA	NA	NA
+ social class [b]	1.1	(1.04-1.17)	1.11	(1.01-1.22)	1.13	(1.02-1.25)	1.26	(1.02-1.55)	1	(0.84-1.20)
+ deprivation category [c]	1.08	(1.02-1.15)	1.1	(1.00-1.22)	1.12	(1.00-1.25)	1.3	(1.04-1.63)	0.99	(0.81-1.21)
Collaborative										
black smoke only	1.06	(1.01-1.11)	1.06	(0.98-1.15)	1.06	(0.97-1.16)	1.09	(0.90-1.32)	1.22	(1.06-1.42)
+ marital status	1.05	(1.00-1.10)	1.06	(0.98-1.14)	1.05	(0.97-1.15)	1.06	(0.87-1.28)	1.19	(1.03-1.38)
+ body mass index	1.04	(0.99-1.09)	1.06	(0.98-1.14)	1.05	(0.96-1.15)	1.04	(0.86-1.27)	1.18	(1.02-1.37)
+ smoking	1.02	(0.97-1.07)	1.03	(0.95-1.11)	1.03	(0.94-1.12)	0.99	(0.81-1.20)	1.12	(0.97-1.30)
+ cholesterol	1.02	(0.97-1.07)	1.04	(0.96-1.12)	1.04	(0.95-1.13)	NA	NA	NA	NA
+ systolic blood pressure	1.02	(0.97-1.07)	1.04	(0.96-1.12)	1.04	(0.95-1.13)	NA	NA	NA	NA
+ social class [b]	1.01	(0.96-1.06)	1.03	(0.95-1.12)	1.03	(0.94-1.12)	0.97	(0.79-1.18)	1.11	(0.96-1.30)
+ deprivation category [c]	1.01	(0.95-1.06)	1.02	(0.93-1.10)	1.01	(0.92-1.11)	0.96	(0.78-1.18)	1.15	(0.98-1.34)

(a) All models are stratified by 1 year age group and gender.
(b) Model presented in Table 2 A and B
(c) Model presented in Table 2C and D.

Table 4. Summary of black smoke-mortality associations in Midspan, NLCS-Air and PAARC cohort studies; and small-area ecological study in Great Britain.

Study	All Cause	Cardiovascular	Respiratory	Lung cancer
R/P cohort [(a)]	1.10 (1.04-1.17)	1.11 (1.01-1.22)	1.26 (1.02-1.55)	1.00 (0.84-1.20)
R/P cohort + Dep [(b)]	1.08 (1.02-1.15)	1.10 (1.00-1.22)	1.30 (1.04-1.63)	0.99 (0.81-1.21)
NLCS-Air cohort [(c)]	1.05 (1.00-1.11)	1.04 (0.95-1.13)	1.22 (0.99-1.50)	1.03 (0.88-1.20)
PAARC cohort [(d)]	1.07 (1.03-1.25)	1.05 (0.98-1.12)[(d)]	1.05 (0.98-1.12)[(d)]	1.03 (0.92-1.15)
GB small area [(e)]	1.019 (1.018-1.021)	1.020 (1.017-1.022)	1.030 (1.026-1.034)	1.026 (1.021-1.032)
GB small area + Dep [(f)]	1.007 (1.006-1.009)	1.007 (1.004-1.009)	1.019 (1.015-1.023)	1.006 (1.000-1.012)

[(a)] BS- mortality association in Renfrew/Paisley cohort observed in present study, with adjustment for individual-level risk factors, including social class.
[(b)] As for [(a)] with adjustment for individual-level risk factors including social class; and additional adjustment area-level deprivation.
[(c)] BS- mortality association in NLCS-Air cohort in Netherlands, including adjustment for area-level socio-economic status [7].
[(d)] BS- mortality association in PAARC cohort in France [5]. Cardiovascular and respiratory BS-mortality associations are not reported separately for PAARC cohort. Therefore the reported BS-'cardiopulmonary' association has been replicated in third and fourth column of above table.
[(e)] BS- mortality association (for 0-8 year exposure window) in small-area ecological study across electoral wards in Great Britain [27].
[(f)] As for [(e)] with additional adjustment for area-level deprivation.

Figure 1: Locations of cohort participants' residential addresses and black smoke monitoring sites.